A NEW CHARACTERIZATION OF QUADRATIC TRANSPORTATION-INFORMATION INEQUALITIES

YUAN LIU

Abstract. It is known that a quadratic transportation-information inequality $\mathrm{W_2I}$ interpolates between the Talagrand's inequality $\mathrm{W_2H}$ and the log-Sobolev inequality (LSI for short). The aim of this paper is threefold:

(1) To prove the equivalence of $\mathrm{W_2I}$ and the Lyapunov condition, which gives a new characterization inspired by Cattiaux-Guillin-Wu [8].
(2) To prove the stability of $\mathrm{W_2I}$ under bounded perturbations, which gives a transference principle in the sense of Holley-Stroock.
(3) To prove $\mathrm{W_2H}$ through a restricted $\mathrm{W_2I}$, which gives a counterpart of the restricted LSI presented by Gozlan-Roberto-Samson [15].

1. Introduction

Transport inequalities have been a very active family of functional inequalities for years, with their profound connections to measure concentration phenomenon and large deviation principle of Markov processes. We refer to two monographies by Villani [25, 26] and two surveys by Gozlan-Léonard [14] and Cattiaux-Guillin [6] on this subject with references therein. In this paper, we will investigate three questions about the quadratic transportation-information inequality $\mathrm{W_2I}$, which interpolates between the Talagrand's inequality $\mathrm{W_2H}$ and the log-Sobolev inequality (LSI for short).

Let's address some basics around the above objects. Denote by (E, d) a metric space equipped with the collection $\mathcal{P}(E)$ of all probability measures on the Borel σ-field. Define the L^p Wasserstein (transportation) distance between two probability measures $\nu, \mu \in \mathcal{P}(E)$, for any $p \geqslant 1$, by

$$W_p(\nu,\mu) = \left(\inf_{\pi \in \mathcal{C}(\nu,\mu)} \int_{E\times E} d^p(x,y)\pi(\mathrm{d}x,\mathrm{d}y) \right)^{1/p},$$

where $\mathcal{C}(\nu, \mu)$ denotes the set of couplings between ν and μ, that is to say the set of probability measures π on $E \times E$ with marginals ν and μ. For simplicity, our framework is specified as the following. Take E to be a connected complete Riemannian manifold of finite dimension, d the geodesic distance, $\mathrm{d}x$ the volume measure, $\mu(\mathrm{d}x) = e^{-V(x)}\mathrm{d}x \in \mathcal{P}(E)$ with $V \in C^1(E)$, $\mathrm{L} = \Delta - \nabla V \cdot \nabla$ the μ-symmetric diffusion operator with domain $\mathbb{D}(\mathrm{L})$, $\Gamma(f, g) = \nabla f \cdot \nabla g$ the carré du champ operator and $\mathcal{E}$ the Dirichlet form with domain $\mathbb{D}(\mathcal{E})$. Denote $\mu(h) := \int h\mathrm{d}\mu$.

Date: June 8, 2016.
2010 *Mathematics Subject Classification.* 26D10, 60E15, 60J60.
Key words and phrases. transportation-information inequality, Talagrand's inequality, log-Sobolev inequality, Lyapunov condition, transference principle.

It's known that the integration by parts formula reads

$$\int \nabla f \cdot \nabla g \, \mathrm{d}\mu = -\int f \mathrm{L} g \, \mathrm{d}\mu, \ \forall f \in \mathbb{D}(\mathcal{E}), g \in \mathbb{D}(\mathrm{L}).$$

The reader is referred to Bakry-Gentil-Ledoux [3] for a detailed presentation.

In the sequel, we focus on the case $p = 2$ of particular interest. Twenty years ago, Talagrand [24] introduced the celebrated transportation-entropy inequality for any ν with respect to the Gaussian measure μ (hence, it is called the Talagrand's inequality)

$$(\mathrm{W_2H(C)}) \qquad W_2(\nu,\mu) \leqslant \sqrt{2CH(\nu|\mu)},$$

where $H(\nu|\mu)$ denotes the relative entropy equal to either $\nu(\log \frac{\mathrm{d}\nu}{\mathrm{d}\mu})$ if ν is absolutely continuous to μ or ∞ otherwise. For general $\mathrm{W_2H}$, various characterizations have been found out, such as the Bobkov-Götze's infimum-convolution criterion (see [5]), the Gozlan-Léonard's large deviation and concentration criteria (see [13]), etc.

As a counterpart, the transportation-information inequality was introduced much later by Guillin-Léonard-Wu-Yao [18], which substituted Fisher-Donsker-Varadhan information for relative entropy, i.e.

$$(\mathrm{W_2I(C)}) \qquad W_2(\nu,\mu) \leqslant \sqrt{2CI(\nu|\mu)},$$

where $I(\nu|\mu)$ is associated to some Dirichlet form $\mathcal{E}$ with domain $\mathbb{D}(\mathcal{E})$ in $L^2(\mu)$ as

$$I(\nu|\mu) = \begin{cases} \mathcal{E}(\sqrt{f},\sqrt{f}), & \text{if } f = \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \text{ with } \sqrt{f} \in \mathbb{D}(\mathcal{E}); \\ \infty, & \text{otherwise.} \end{cases}$$

Some criteria for $\mathrm{W_2I}$ have also been worked out correspondingly by [18].

The reference measure μ is usually regarded as the unique invariant distribution of (symmetric) diffusion process. However in some practical cases, μ is unknown except its existence. For this reason, it is natural to present suitable hypotheses on the infinitesimal generator instead of its equilibrium limit. Going to this direction, Cattiaux-Guillin-Wang-Wu [7] drew in the Lyapunov condition to study the super Poincaré inequalities. Afterwards, Cattiaux-Guillin-Wu [8] derived $\mathrm{W_2H}$ from the Lyapunov condition, which even worked for LSI

$$H(\nu|\mu) \leqslant 2CI(\nu|\mu)$$

with additional assumptions on the Bakry-Emery's curvature.

More precisely as [8], say $W \geqslant 1$ is a Lyapunov function if there exist two constants $c > 0, b \geqslant 0$ and some $x_0 \in E$ such that in the sense of distribution

$$\mathrm{L}W \leqslant \left(-cd^2(x,x_0) + b\right) W. \tag{1.1}$$

There were also some variants of (1.1) by substituting $d^2(\cdot,\cdot)$ with other functionals of distance, such as [2, 9] for investigating relations between the Poincaré inequalities and weak Lyapunov conditions.

Theorem 1.1. *(*[8, Theorem 1.2]*) Under the Lyapunov condition (1.1),*

1. *There exists a constant $C_1 > 0$ such that μ verifies the inequality* $\mathrm{W_2H}(C_1)$.
2. *There exists a constant $C_2 > 0$ such that μ verifies the inequality* $\mathrm{LSI}(C_2)$ *provided that the Bakry-Emery's curvature is bounded from below.*

In the preprint [21], the author tried to show that if the Bakry-Emery's curvature has a lower bound, the LSI and Lyapunov condition (with a slight adjustment) are equivalent. On the other hand, due to the well-known HWI inequality from Otto-Villani [23], it follows that LSI and $\mathrm{W_2I}$ are also equivalent under the same curvature condition. These results lead to our first question.

Question 1. *Is it possible to get* $\mathrm{W_2I}$ *through the Lyapunov condition without any assumption on the curvature? And what about the converse implication?*

For some technical reasons, certain regularity condition is usually imposed on the density of μ. One way of removing such a constraint is to handle a nice model space firstly and then extend the associated functional inequalities to a general setting through perturbations. For example, Holley-Stroock [19] proved that if μ satisfies a LSI and $\tilde{\mu}$ is a bounded variant of μ with $\frac{\mathrm{d}\tilde{\mu}}{\mathrm{d}\mu}$ and $\frac{\mathrm{d}\mu}{\mathrm{d}\tilde{\mu}} \in L^\infty$, then $\tilde{\mu}$ verifies a LSI too. This is called the stability or transference principle of inequalities.

For LSI, perturbations on the entropy and information can be respectively and directly controlled. Somehow it is tough to deal with the Wasserstein distance, so that the stability of $\mathrm{W_2H}$ remained open until it was attacked by Gozlan-Roberto-Samson [15] via a new characterization, the so called restricted LSI (rLSI for short). They defined f to be a K-semi-convex function if for any $x, y \in \mathbb{R}^n$ (which can be extended to manifolds or length spaces, etc.)

$$f(y) \geqslant f(x) + \nabla f(x) \cdot (y - x) - \frac{K}{2} d^2(x, y),$$

see also the classical semi-convexity in the textbook of Evans [11, Section 3.3]. Say μ verifies a rLSI with constant $C > 0$ if for all K-semi-convex f with $0 \leqslant K < C^{-1}$

$$\text{(rLSI(C))} \qquad \mathrm{Ent}_\mu(e^f) \leqslant \frac{2C}{(1 - KC)^2} \int |\nabla f|^2 e^f \mathrm{d}\mu,$$

where $\mathrm{Ent}_\mu(e^f) = \mu(fe^f) - \mu(e^f)\log\mu(e^f)$, and actually $\int |\nabla f|^2 e^f \mathrm{d}\mu = 4I(e^f)$. We quote partial results from [15] as follows.

Theorem 1.2. *([15, Theorem 1.5]) The next two statements are equivalent:*

(1) *There exists a constant $C_1 > 0$ such that μ verifies the inequality* $\mathrm{W_2H}(C_1)$.
(2) *There exists a constant $C_2 > 0$ such that μ verifies the inequality* $\mathrm{rLSI(C_2)}$.

A quantitative relationship between control constants is the following: from (1) to (2) holds $C_1 = C_2$, and conversely, from (2) to (1) holds $C_1 = 8C_2$.

Now our second question arises.

Question 2. *What about the stability of* $\mathrm{W_2I}$ *in the sense of Holley-Stroock?*

It is known that $\mathrm{W_2I}$ implies $\mathrm{W_2H}$ according to Guillin-Léonard-Wang-Wu [17, Theorem 2.4], but it is not clear yet to what extent they are different. Moerover, $\mathrm{W_2I}$ is an interpolation between $\mathrm{W_2H}$ and LSI, which suggests that this relationship should be kept for the restricted type inequalities. So our third question comes out.

Question 3. *What is the restricted* $\mathrm{W_2I}$ *equivalent to* $\mathrm{W_2H}$*?*

This paper will answer the above three questions. With a slight adjustment to (1.1), say $W > 0$ is a Lyapunov function if W^{-1} is locally bounded and there exist two constants $c > 0, b \geqslant 0$ and some $x_0 \in E$ such that in the sense of distribution

$$\text{(1.2)} \qquad \mathrm{L}W \leqslant \left(-cd^2(x, x_0) + b\right) W.$$

Here we use $W > 0$ instead of $W \geqslant 1$ in (1.1), but the technique in the proof of Bakry-Barthe-Cattiaux-Guillin [1, Theorem 1.4] still works if W^{-1} is locally bounded (not necessary to request a uniform lower bound). We will go back to this point in Section 2.

What we prove for the first question is the following.

Theorem 1.3. *The next two statements are equivalent:*

1. *There exists a constant $C > 0$ such that μ verifies the inequality* $\mathrm{W_2I}(C)$.
2. *There exists a Lyapunov function $W > 0$ that verifies Condition (1.2).*

On the second question, our method is not to look for an equivalent restricted type LSI for $\mathrm{W_2I}$, but turns to applying Theorem 1.3 and 1.2 directly.

Theorem 1.4. *Let $\tilde{\mu} \in \mathcal{P}(E)$ be absolutely continuous to μ with $M^{-1} \leqslant \frac{\mathrm{d}\tilde{\mu}}{\mathrm{d}\mu} \leqslant M$ for some constant $M \geqslant 1$. Then $\tilde{\mu}$ verifies the inequality* $\mathrm{W_2I}$ *if so does μ.*

Similarly to [15], we introduce if $\log \frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ is K-semi-convex with $0 \leqslant K < C^{-1}$

$$W_2(\nu, \mu) \leqslant \sqrt{\frac{4C^2}{(1-KC)^2} I(\nu|\mu)}. \tag{rW$_2$I(C)}$$

The above constant is chosen for convenience. For the third question, we prove

Theorem 1.5. *The next two statements are equivalent:*

1. *There exists a constant $C_1 > 0$ such that μ verifies the inequality* $\mathrm{W_2H}(C_1)$.
2. *There exists a constant $C_2 > 0$ such that μ verifies the inequality* $\mathrm{rW_2I}(C_2)$.

A quantitative relationship between control constants is the following: from (1) *to* (2) *holds $2C_1 = C_2$, and conversely, from* (2) *to* (1) *holds $C_1 = 4C_2$.*

The rest of this paper contains three sections, which give the proofs successively for Theorems 1.3, 1.4 and 1.5. An alternative proof of $\mathrm{W_2H}$ through the Lyapunov condition (1.2) is also provided, see Section 4.

2. Equivalence of W$_2$I and the Lyapunov condition

Throughout this paper, write $d_0^2(x) := d^2(x, x_0)$. The use of Lyapunov condition (1.2) is based on [1, Theorem 1.4].

Lemma 2.1. *Under (1.2), for any $h \in \mathbb{D}(\mathcal{E})$*

$$\int h^2 d_0^2 \mathrm{d}\mu \leqslant \frac{1}{c}\int |\nabla h|^2 \mathrm{d}\mu + \frac{b}{c}\int h^2 \mathrm{d}\mu. \tag{2.1}$$

Moreover, the Poincaré inequality holds, i.e. there exists $C > 0$ such that

$$\int |h - \mu(h)|^2 \mathrm{d}\mu \leqslant C \int |\nabla h|^2 \mathrm{d}\mu. \tag{PI(C)}$$

Proof. The technique in [1, Page 64] yields

$$\begin{aligned}
\int h^2 d_0^2 \mathrm{d}\mu &= \frac{1}{c}\int h^2 (c d_0^2 - b)\mathrm{d}\mu + \frac{b}{c}\int h^2 \mathrm{d}\mu \\
&\leqslant \frac{1}{c}\int \frac{-\mathrm{L}W}{W} h^2 \mathrm{d}\mu + \frac{b}{c}\int h^2 \mathrm{d}\mu \\
&= \frac{1}{c}\int \nabla W \cdot \nabla \frac{h^2}{W} \mathrm{d}\mu + \frac{b}{c}\int h^2 \mathrm{d}\mu
\end{aligned}$$

$$\begin{aligned} &= \frac{1}{c}\int |\nabla h|^2 - \left|\nabla h - \frac{h}{W}\nabla W\right|^2 \mathrm{d}\mu + \frac{b}{c}\int h^2 \mathrm{d}\mu \\ &\leqslant \frac{1}{c}\int |\nabla h|^2 \mathrm{d}\mu + \frac{b}{c}\int h^2 \mathrm{d}\mu. \end{aligned}$$

Note that here is no need to assume the integrability of d_0^2 or $\frac{-LW}{W}$ for μ, since we can take an approximation sequence in $C_c^\infty(E)$ for given h.

Next, according to [1], let B be a ball of radius r centered at x_0 with $r^2 = \frac{b}{c}+1$. Denote $\bar{h}_B = \mu(h\mathbf{1}_B)/\mu(B)$, we have

$$\begin{aligned} \int |h-\mu(h)|^2 \mathrm{d}\mu &= \inf_{\lambda\in\mathbb{R}} \int |h-\lambda|^2 \mathrm{d}\mu \;\leqslant\; \int |h-\bar{h}_B|^2 \mathrm{d}\mu \\ &= \int |h-\bar{h}_B|^2 d_0^2 \mathrm{d}\mu + \int |h-\bar{h}_B|^2 (1-d_0^2)\mathrm{d}\mu \\ &\leqslant \frac{1}{c}\int |\nabla h|^2 \mathrm{d}\mu + \left(\frac{b}{c}+1\right)\int_B |h-\bar{h}_B|^2 \mathrm{d}\mu, \end{aligned}$$

which implies the (global) Poincaré inequality by combining the local one on B. □

Remark 2.2. *Under (1.1), it was proved that μ admits the Gaussian integrability $\mu e^{\delta d_0^2} < \infty$ for some $\delta > 0$. In* [20], *an elementary proof of this fact was given with a sharp estimate for δ. Actually, these all hold under (1.2) too.*

Recall Barthe-Cattiaux-Roberto [4, Lemma 14], the following inequality holds.

Lemma 2.3. *For any $s \in [0, Nt]$ with $N \geqslant 0$ and $t > 0$,*

$$s^2 \log \frac{s^2}{t^2} - (s^2 - t^2) \leqslant (1+N)^2 (s-t)^2.$$

Proof. Use the Lagrange's mean-value theorem to get $\log a - \log b \leqslant \frac{a-b}{b}$, and then substitute $\frac{s^2}{t^2}$ for a and 1 for b. □

Recall the definition of infimum-convolution $Q_t h$ is for any $t \geqslant 0$

$$Q_t h(x) := \inf_{y\in E}\left\{h(y) + \frac{1}{2}d^2(x,y)\right\}.$$

Lemma 2.4. *Let $g = \frac{1}{2}(d_0^2 \wedge D)$ for any constant $D \in [0,\infty]$. Then $Q_1 g \geqslant \frac{1}{2}g$.*

Proof. The definition gives

$$Q_1 g(x) = \inf_y \left\{g(y) + \frac{1}{2}d^2(x,y)\right\} = \frac{1}{2}\inf_y \left\{d^2(y,x_0)\wedge D + d^2(x,y)\right\}.$$

Suppose the infimum is achieved at some z, it follows either $d^2(z,x_0) < D$ so that

$$d^2(z,x_0)\wedge D + d^2(x,z) = d^2(z,x_0) + d^2(x,z) \geqslant \frac{1}{2}d^2(x,x_0) \geqslant \frac{1}{2}\left(d^2(x,x_0)\wedge D\right),$$

or $d^2(z,x_0) \geqslant D$ so that

$$d^2(z,x_0)\wedge D + d^2(x,z) = D + d^2(x,z) \geqslant D \geqslant d^2(x,x_0)\wedge D.$$

Hence, it is always true that

$$Q_1 g(x) \geqslant \frac{1}{4}\left(d^2(x,x_0)\wedge D\right) \;=\; \frac{1}{2}g(x).$$

The proof is completed. □

Now we prove Theorem 1.3.

Proof. The strategy contains two parts.

Part 1. Assume the Lyapunov condition (1.2) holds. Then $\mathrm{W_2H}(C_T)$ comes true with constant $C_T > 0$ by [8, Theorem 1.9], which also verifies the Bobkov-Götze's infimum-convolution criterion in [5] that for any bounded h with $\mu(h) = 0$,

$$\int e^{Q_{C_T}h}\mathrm{d}\mu \leqslant 1. \tag{2.2}$$

From the representation (see for example Bakry-Gentil-Ledoux [3, Section 9.2])

$$\frac{1}{2}W_2(\nu,\mu)^2 = \sup_{h \text{ bounded}} \int Q_1 h\mathrm{d}\nu - \int h\mathrm{d}\mu = \sup_{\substack{h \text{ bounded} \\ \mu(h)=0}} \int Q_1 h\mathrm{d}\nu,$$

it follows

$$\frac{1}{2C_T}W_2(\nu,\mu)^2 = \sup_{\substack{h \text{ bounded} \\ \mu(h)=0}} \int Q_{C_T} h\mathrm{d}\nu.$$

Let $\mathrm{d}\nu = f^2\mathrm{d}\mu$ with $f > 0$. We introduce a subset with some parameter $N > 1$

$$A = \{x : 0 < f \leqslant N \cdot \mu(f)\},$$

to make the following decomposition

$$\int f^2 Q_{C_T} h\mathrm{d}\mu = \int_A f^2(Q_{C_T}h - \log f^2)\mathrm{d}\mu + \int_A f^2 \log f^2 \mathrm{d}\mu + \int_{A^c} f^2 Q_{C_T} h\mathrm{d}\mu.$$

Now we estimate each term in the above sum. First of all, using the inequality $\log a - \log b \leqslant \frac{a-b}{b}$ and (2.2) yields

$$\begin{aligned}
\int_A f^2(Q_{C_T}h - \log f^2)\mathrm{d}\mu &\leqslant \int_A e^{Q_{C_T}h} - f^2\mathrm{d}\mu \\
&= \int e^{Q_{C_T}h} - f^2\mathrm{d}\mu + \int_{A^c} f^2 - e^{Q_{C_T}h}\mathrm{d}\mu \\
&\leqslant \int_{A^c} f^2\mathrm{d}\mu \leqslant \frac{N^2}{(N-1)^2}\int_{A^c} (f-\mu(f))^2\,\mathrm{d}\mu. \qquad (2.3)
\end{aligned}$$

The last inequality is due to $f > N \cdot \mu(f)$ on A^c.

Next, since $0 < (\mu(f))^2 \leqslant \mu(f^2) = 1$, we have by Lemma 2.3 for $t = \mu(f)$ and $s = f(x)$ with $x \in A$ that

$$\begin{aligned}
\int_A f^2 \log f^2\mathrm{d}\mu &= \int_A f^2 \log \frac{f^2}{(\mu(f))^2}\mathrm{d}\mu + \int_A f^2\mathrm{d}\mu \cdot \log (\mu(f))^2 \\
&\leqslant \int_A f^2 - (\mu(f))^2 + (1+N)^2 (f-\mu(f))^2\,\mathrm{d}\mu \\
&\leqslant \int f^2 - (\mu(f))^2\,\mathrm{d}\mu + (1+N)^2\int_A (f-\mu(f))^2\,\mathrm{d}\mu \\
&\leqslant \left[1+(1+N)^2\right]\int (f-\mu(f))^2\,\mathrm{d}\mu. \qquad (2.4)
\end{aligned}$$

Thirdly, the definition of infimum-convolution gives

$$Q_{C_T}h(x) \leqslant \int h(y) + \frac{1}{2C_T}d^2(x,y)\mathrm{d}\mu(y) \leqslant \frac{1}{C_T}\left(d_0^2(x) + \mu(d_0^2)\right),$$

and then due to Lemma 2.1

$$\begin{aligned}\int_{A^c} f^2 Q_{C_T} h \,\mathrm{d}\mu &\leqslant \frac{N^2}{C_T(N-1)^2}\int_{A^c}(f-\mu(f))^2\left(d_0^2+\mu(d_0^2)\right)\mathrm{d}\mu\\ &\leqslant \frac{N^2}{C_T(N-1)^2}\left(\frac{1}{c}\int|\nabla f|^2\mathrm{d}\mu+\frac{b+\mu(d_0^2)}{c}\int(f-\mu(f))^2\,\mathrm{d}\mu\right).\end{aligned}\tag{2.5}$$

Combining (2.3-2.5) and using the Poincaré inequality PI(C_P) yield

$$W_2(\nu,\mu)\leqslant\sqrt{2CI(\nu|\mu)},$$

where C is less than $C_TC_P[\frac{N^2}{(N-1)^2}+1+(1+N)^2]+\frac{N^2}{c(N-1)^2}\left[1+\left(b+\mu(d_0^2)\right)C_P\right]$.

Part 2. Assume $\mathrm{W_2I}(C)$ holds, which implies $\mathrm{W_2H}(C)$ by [18, Theorem 2.4] and then $\mathrm{W_1H}(C)$ automatically. According to Djellout-Guillin-Wu [10, Theorem 2.3], μ satisfies the Gaussian integrability $\mu(e^{\delta d_0^2})<\infty$ for some $\delta>0$.

A Lyapunov function can be constructed by solving certain elliptic equation. Set $\phi=-cd_0^2+b$ with two parameters $c>0, b\geqslant 0$, which will be determined below. Introduce a partial differential equation of second order for $w\in L^2(\mu)$

$$\mathrm{H}u := -\mathrm{L}u+\phi u = w.\tag{2.6}$$

First of all, (2.6) gives $\mu(u\cdot\mathrm{H}u)\leqslant\mathcal{E}[u]+b\mu(u^2)$ directly. On the other hand, for any $u\in\mathbb{D}(\mathcal{E})$ with $\mu(u^2)>0$, let $\mathrm{d}\nu=u^2/\mu(u^2)\mathrm{d}\mu$ and $g_D=\frac{1}{2}(d_0^2\wedge D)$, using Lemma 2.4 and $\mathrm{W_2I}$ yields

$$\begin{aligned}\mu\left(u^2d_0^2\right) &= \lim_{D\to\infty}\mu\left(u^2(d_0^2\wedge D)\right)\leqslant\lim_{D\to\infty}4\mu\left(u^2Q_1g_D\right)\\ &\leqslant \lim_{D\to\infty}4\mu(u^2)W_2(\nu,\mu)^2+4\mu(u^2)\mu(g_D)\leqslant 8C\mathcal{E}[u]+2\mu(u^2)\mu(d_0^2),\end{aligned}$$

which implies by taking $c=1/(16C)$ and $b=4c\cdot\mu(d_0^2)$ that

$$\int cu^2d_0^2\mathrm{d}\mu\leqslant\frac{1}{2}\left(\mathcal{E}[u]+b\mu(u^2)\right),\tag{2.7}$$

and then

$$\begin{aligned}\mu(u\cdot\mathrm{H}u) &= \mathcal{E}[u]+b\mu(u^2)-c\mu\left(u^2d_0^2\right)\\ &\geqslant \mathcal{E}[u]+b\mu(u^2)-c\left(8C\mathcal{E}[u]+2\mu(u^2)\mu(d_0^2)\right)\geqslant\frac{1}{2}\left(\mathcal{E}[u]+b\mu(u^2)\right).\end{aligned}$$

Combining these estimates with the Hölder inequality yields for any $u\in\mathbb{D}(\mathcal{E})$ and $v\in\mathbb{D}(\mathrm{L})$

$$\begin{aligned}(u,\mathrm{H}v) &= \int\nabla u\cdot\nabla v-cd_0^2uv+buv\,\mathrm{d}\mu\\ &\leqslant (\mathcal{E}[u])^{\frac12}(\mathcal{E}[v])^{\frac12}+c\left(\mu(d_0^2u^2)\right)^{\frac12}\left(\mu(d_0^2v^2)\right)^{\frac12}+b\left(\mu(u^2)\right)^{\frac12}\left(\mu(v^2)\right)^{\frac12}\\ &\leqslant \frac{5}{2}\left(\mathcal{E}[u]+b\mu(u^2)\right)^{\frac12}\left(\mathcal{E}[v]+b\mu(v^2)\right)^{\frac12}.\end{aligned}$$

Hence, $(u,\mathrm{H}v)$ determines a coercive Dirichlet form, and H is a positive definite self-adjoint Schrödinger operator with its spectrum contained in $(0,\infty)$. It means H^{-1} exists on $L^2(\mu)$ according to the Lax-Milgram Theorem, i.e. $u=\mathrm{H}^{-1}w\in H^1(\mu)$ (the L^2-integrable Sobolev space of weak derivatives of first order) is a weak solution of Equation (2.6), see Evans [11] or Gilbarg-Trudinger [12].

Whenever $w \geqslant 0$, the weak maximum principle yields $u = \mathrm{H}^{-1}w \geqslant 0$ μ-a.e. too. As a routine, we set $u_- = -\min\{u, 0\}$, which has weak derivatives and satisfies

$$\begin{aligned} 0 \leqslant \mu(u_- w) &= \mu(u_- \cdot \mathrm{H}u) = \int u_- \cdot \mathrm{L}u\mathrm{d}\mu - \int \phi u_- u\mathrm{d}\mu \\ &= -\int \nabla u_- \cdot \nabla u\mathrm{d}\mu - \mu(\phi u_-^2) \\ &= -\mathcal{E}[u_-] - b\mu(u_-^2) + c\mu(d_0^2 u_-^2) \leqslant -\frac{1}{2}\left(\mathcal{E}[u_-] + b\mu(u_-^2)\right) \leqslant 0, \end{aligned}$$

where the middle inequality is derived from (2.7). It follows $\mu(u_-^2) = 0$, which implies $u_- = 0$ μ-a.e. and then $u \geqslant 0$ μ-a.e.

Now fix $w \equiv 1$ and $u = \mathrm{H}^{-1}1$. For any ball $B \subset E$, $u\mathbf{1}_B$ gives a weak solution to Equation (2.6) restricted in B, since there holds for any $h \in C_{\mathrm{c}}^1(B)$ (i.e. the set of first-order derivative functions with compact support in B)

$$\int_B \mathrm{H}u \cdot h\mathrm{d}x = \int \mathrm{H}u \cdot (he^V)\mathrm{d}\mu = \int w \cdot (he^V)\mathrm{d}\mu = \int_B w \cdot h\mathrm{d}x.$$

When E is the Euclidean space, according to [12, Theorem 8.22] and the notation therein, u is locally Hölder continuous in B if we set $f^i = 0$, $g = -w$ and $L = \mathrm{L} - \phi$ such that $Lu = g$ as [12] did. Note that the continuity is a local property. Since any local region in Riemannian manifold is (smoothly) diffeomorphic to a region in $\mathbb{R}^n$, which preserves the uniform ellipticity for the (weighted) Laplacian, it follows that u is continuous on E in the framework of Riemannian manifolds.

Moreover, we prove that $u > 0$ everywhere. By contradiction, assume $u(y) = 0$ at some y. Choose $r > 0$ and $v \in C^2(E)$ to satisfy

$$v \geqslant 0,\ v(y) > 0,\ v_{|B_r(y)^c} = 0,\ \mathrm{H}v \leqslant \frac{1}{2}.$$

It follows on E

$$\mathrm{H}(u - v) = w - \mathrm{H}v \geqslant \frac{1}{2},$$

which implies $u - v \geqslant 0$ μ-a.e. by the weak maximum principle. Then $u(y) \geqslant v(y)$ by the continuity, which is absurd.

As consequence, we obtain $u \in H^1(\mu) \cap C(E)$ is strictly positive everywhere, and u^{-1} is locally bounded. So u is a Lyapunov function for $\mathrm{L}u \leqslant \phi u$. $\square$

3. Stability of $\mathrm{W_2I}$

Now we prove Theorem 1.4, with similar arguments for Theorem 1.3 and a few more efforts.

Proof. Assume μ verifies $\mathrm{W_2I}(C)$. Thanks to the implication from $\mathrm{W_2I}$ to $\mathrm{W_2H}$ by [18] and the stability of $\mathrm{W_2H}$ by [15], $\tilde{\mu}$ also verifies the Bobkov-Götze's criterion, namely there exists $\tilde{C_T} > 0$ such that for any bounded h with $\tilde{\mu}(h) = 0$,

$$\int e^{Q_{\tilde{C_T}} h}\mathrm{d}\tilde{\mu} \leqslant 1.$$

Recall the first part of proof for Theorem 1.3, we introduce $\mathrm{d}\tilde{\nu} = f^2\mathrm{d}\tilde{\mu}$ with $f > 0$ and $\tilde{A} = \{x : 0 < f \leqslant N \cdot \tilde{\mu}(f)\}$, and then deal with $\int f^2 Q_{\tilde{C_T}} h\mathrm{d}\tilde{\mu}$ by combining three estimates similar as (2.3-2.5).

But there is a gap since (2.5) relies on the Lyapunov condition, which holds for μ by Theorem 1.3, not proved for $\tilde{\mu}$ yet. Nevertheless, it can be quickly fixed due to $\frac{\mathrm{d}\tilde{\mu}}{\mathrm{d}\mu}$ is two-sided bounded so that

$$\begin{aligned}\int_{\tilde{A}^c} f^2 Q_{\tilde{C_T}} h \,\mathrm{d}\tilde{\mu} &\leqslant \frac{N^2}{\tilde{C_T}(N-1)^2}\int_{\tilde{A}^c} (f-\tilde{\mu}(f))^2 \left(d_0^2+\mu(d_0^2)\right)\mathrm{d}\tilde{\mu} \\ &\leqslant \frac{MN^2}{\tilde{C_T}(N-1)^2}\int_{\tilde{A}^c} (f-\tilde{\mu}(f))^2 \left(d_0^2+\mu(d_0^2)\right)\mathrm{d}\mu \\ &\leqslant \frac{MN^2}{\tilde{C_T}(N-1)^2}\left(\frac{1}{c}\int |\nabla f|^2 \mathrm{d}\mu + \frac{b+\mu(d_0^2)}{c}\int (f-\tilde{\mu}(f))^2\,\mathrm{d}\mu\right) \\ &\leqslant \frac{M^2N^2}{\tilde{C_T}(N-1)^2}\left(\frac{1}{c}\int |\nabla f|^2 \mathrm{d}\tilde{\mu} + \frac{b+\mu(d_0^2)}{c}\int (f-\tilde{\mu}(f))^2\,\mathrm{d}\tilde{\mu}\right). \qquad (3.1)\end{aligned}$$

Hence, we can substitute (3.1) for (2.5) and then complete the proof by using the stability of Poincaré inequality. □

4. A characterization of $\mathrm{W_2H}$ via a restricted $\mathrm{W_2I}$

This section has two parts. First, we give a direct proof $\mathrm{W_2H}$ under the Lyapunov condition, which was originally proposed by [8, Theorem 1.2]. Next, we introduce a restricted $\mathrm{W_2I}$ interpolating between $\mathrm{W_2H}$ and the restricted LSI in the sense of [15, Definition 1.3]. We would like to point out here, the theory of Hamilton-Jacobi equations played a fundamental role in studying quadratic transport inequalities. The general work by Gozlan-Roberto-Samson [16] on metric spaces, which extended some early results by Lott-Villani [22], is sufficiently adapted in the framework of Riemannian manifolds.

4.1. **An alternative proof of** $\mathrm{W_2H}$ **under the Lyapunov condition (1.2).** Now we prove Theorem 1.1.

Proof. Let $\delta > 0$ be some parameter. Given any bounded h with $\mu(h) = 0$, define for all $x \in E$ and $t > 0$

$$\phi(x,t) = \delta t Q_t h(x), \quad \Lambda = \mu(e^{\phi}), \quad \lambda = \mu(e^{\phi/2}).$$

According to [16], the Hopf-Lax formula $Q_t h$ solves the Hamilton-Jacobi equation

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}u + \frac{1}{2}|\nabla u|^2 = 0, \\ u_0 = h, \end{cases}$$

and we have

$$\frac{\mathrm{d}\Lambda}{\mathrm{d}t} = \int e^{\phi}\left(\delta Q_t h + \delta t \frac{\mathrm{d}Q_t h}{\mathrm{d}t}\right)\mathrm{d}\mu = \frac{1}{\delta t}\int \delta e^{\phi}\phi - 2|\nabla e^{\phi/2}|^2 \mathrm{d}\mu. \qquad (4.1)$$

We need to estimate the following

$$\int e^{\phi}\phi \mathrm{d}\mu = \int e^{\phi}\log\frac{e^{\phi}}{\lambda^2}\mathrm{d}\mu + \Lambda\log\lambda^2.$$

Set $A = \{x : e^{\phi/2} \leqslant N\lambda\}$ for any given $N > 1$, which implies $\lambda^2 < e^{\phi}$ on A^c. Using Lemma 2.3 on A yields

$$\int_A e^{\phi}\log\frac{e^{\phi}}{\lambda^2}\mathrm{d}\mu \leqslant \int_A e^{\phi} - \lambda^2 + (1+N)^2(e^{\phi/2}-\lambda)^2\mathrm{d}\mu$$

$$
\begin{aligned}
&= \int e^{\phi} - \lambda^2 \mathrm{d}\mu + \int_{A^c} \lambda^2 - e^{\phi}\mathrm{d}\mu + (1+N)^2 \int_A (e^{\phi/2}-\lambda)^2 \mathrm{d}\mu \\
&= \int (e^{\phi/2}-\lambda)^2 \mathrm{d}\mu + \int_{A^c} \lambda^2 - e^{\phi}\mathrm{d}\mu + (1+N)^2 \int_A (e^{\phi/2}-\lambda)^2 \mathrm{d}\mu \\
&\leqslant \left[1+(1+N)^2\right] \mu(e^{\phi/2}-\lambda)^2. \qquad (4.2)
\end{aligned}
$$

Next, let $d_\mu^2(x) = \mu\left(d^2(x,\cdot)\right)$, we have due to the facts $e^{\phi} \leqslant \frac{N^2}{(N-1)^2}(e^{\phi/2}-\lambda)^2$ on A^c and $\phi \leqslant \frac{\delta}{2} d_\mu^2$ on E

$$
\begin{aligned}
\int_{A^c} e^{\phi} \log \frac{e^{\phi}}{\lambda^2} \mathrm{d}\mu &= \int_{A^c} e^{\phi}\phi \mathrm{d}\mu - \int_{A^c} e^{\phi} \mathrm{d}\mu \log \lambda^2 \\
&\leqslant \frac{\delta N^2}{2(N-1)^2} \int_{A^c} (e^{\phi/2}-\lambda)^2 d_\mu^2 \mathrm{d}\mu - \int_{A^c} e^{\phi}\mathrm{d}\mu \log \lambda^2. \qquad (4.3)
\end{aligned}
$$

It follows from Lemma 2.1 and $d_\mu^2 \leqslant 2d_0^2 + 2\mu(d_0^2)$ that

$$
\int_{A^c} (e^{\phi/2}-\lambda)^2 d_\mu^2 \mathrm{d}\mu \leqslant \frac{2}{c} \int |\nabla e^{\phi/2}|^2 \mathrm{d}\mu + \frac{2b + 2\mu(d_0^2)}{c} \int (e^{\phi/2}-\lambda)^2 \mathrm{d}\mu. \qquad (4.4)
$$

Combining (4.2-4.4) with the Poincaré inequality PI(C_P) gives

$$
\int e^{\phi}\phi \mathrm{d}\mu \leqslant C_0 \int |\nabla e^{\phi/2}|^2 \mathrm{d}\mu + \int_A e^{\phi}\mathrm{d}\mu \log \lambda^2, \qquad (4.5)
$$

where $C_0 = C_P[1+(1+N)^2] + \frac{\delta N^2\left[1+\left(b+\mu(d_0^2)\right)C_P\right]}{c(N-1)^2}$.

Combining (4.1) and (4.5), we take $\delta > 0$ with $\delta C_0 = 2$ so that

$$
\frac{\mathrm{d}\Lambda}{\mathrm{d}t} \leqslant \frac{1}{t} \int_A e^{\phi}\mathrm{d}\mu \log \lambda^2 \leqslant \frac{1}{t}(\Lambda \log \Lambda) \vee 0.
$$

The rest work is similar to the last step in the proof of [8, Lemma 3.2]. Note that $\Lambda(0) = 1$. If $\Lambda(1) > 1$, let $t_0 \in [0,1)$ be the maximal time such that $\Lambda(t_0) = 1$. Then for all $t \in (t_0, 1)$ holds $\frac{\mathrm{d}\Lambda}{\mathrm{d}t} \leqslant \frac{\Lambda \log \Lambda}{t}$, which means $\frac{\mathrm{d}}{\mathrm{d}t}(\frac{\log \Lambda}{t}) \leqslant 0$ and thus $\log \Lambda(1) \leqslant \lim_{t \downarrow t_0} \frac{\log \Lambda(t)}{t} = 0$. It contradicts the assumption $\Lambda(1) > 1$. Hence, the Bobkov-Götze's criterion is verified and $\mathrm{W_2H}(\frac{1}{\delta})$ follows. □

4.2. **A restricted** $\mathrm{W_2I}$. According to [15, 16], define the supremum-convolution

$$
P_t h(x) = \sup_{y \in E} \left\{ h(y) - \frac{1}{2t} d^2(x,y) \right\},
$$

which solves the Hamilton-Jacobi equation

$$
\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} u - \frac{1}{2}|\nabla u|^2 = 0, \\ u_0 = h, \end{cases}
$$

When h is Lipschitz continuous, there is a unique weak solution to the equation (see also Evans [11, Section 3.3, Theorem 7]), which means $P_{1-t} f = Q_t P_1 f$ for $t \in [0,1]$. Moreover, $P_t h$ is t^{-1}-semi-convex by [15, Lemma 5.3].

Now we prove Theorem 1.5.

Proof. If $\mathrm{W_2H}(C_T)$ holds, we have the $\mathrm{rLSI}(C_T)$ for all K-semi-convex functions with $0 \leqslant K < C_T^{-1}$ by [15, Theorem 1.5]. Then for K-semi-convex f with $\mu(e^f) = 1$

$$W_2(e^f\mu, \mu)^2 \leqslant 2C_T \mathrm{Ent}_\mu(e^f) \leqslant \frac{16C_T^2}{(1-KC_T)^2} I(e^f),$$

which gives the $\mathrm{rW_2I}(2C_T)$.

On the other hand, if $\mathrm{rW_2I}(C)$ holds, we adjust the proof of Theorem 1.1 to get $\mathrm{W_2H}$. Let $\delta > 0$, given any bounded Lipschitz function h, we introduce for all $x \in E$ and $t \in [0, \frac{1}{2}]$

$$\psi(x,t) = \delta t P_{1-t}h(x) - \delta t \mu(P_1 h), \quad \Lambda = \mu(e^\psi).$$

It follows due to $P_{1-t}h = Q_t P_1 h$

$$\begin{aligned} \frac{\mathrm{d}\Lambda}{\mathrm{d}t} &= \int e^\psi \left(\delta P_{1-t}h + \delta t \frac{\mathrm{d}P_{1-t}h}{\mathrm{d}t} - \delta\mu(P_1 h)\right)\mathrm{d}\mu \\ &= \frac{1}{\delta t}\int \delta^2 t Q_t \left(P_1 h - \mu(P_1 h)\right) e^\psi - 2|\nabla e^{\psi/2}|^2 \mathrm{d}\mu. \end{aligned} \tag{4.6}$$

Let $g = P_1 h - \mu(P_1 h)$, since ψ is δ-semi-convex, we have by using $\mathrm{rW_2I}(C)$ for any $0 < \delta \leqslant \frac{1}{2C}$

$$\begin{aligned} \int \delta^2 t Q_t g e^\psi \mathrm{d}\mu &= \delta^2 \int Q_1(tg) e^\psi \mathrm{d}\mu \\ &\leqslant \frac{\delta^2 \mu(e^\psi)}{2} W_2\left(\frac{e^\psi}{\mu(e^\psi)}\mu, \mu\right)^2 \\ &\leqslant \frac{2\delta^2 C^2}{(1-\delta C)^2} I(e^\psi) \leqslant 2I(e^\psi). \end{aligned} \tag{4.7}$$

Combining (4.6-4.7) gives $\frac{\mathrm{d}\Lambda}{\mathrm{d}t} \leqslant 0$, which implies due to $P_{1-t}h \geqslant h$

$$\begin{aligned} 1 = \Lambda(0) \geqslant \Lambda\left(\frac{1}{2}\right) &= \mu\left(\exp\left\{\frac{\delta}{2}\left(P_{\frac{1}{2}}h - \mu(P_1 h)\right)\right\}\right) \\ &\geqslant \mu\left(\exp\left\{\frac{\delta}{2}h - \mu\left(P_{\frac{2}{\delta}}\left(\frac{\delta}{2}h\right)\right)\right\}\right). \end{aligned}$$

Hence, we obtain $\mathrm{W_2H}(\frac{2}{\delta})$ by the Bobkov-Götze's supremum-convolution criterion (see [15, Theorem 3.1]). Choosing $\delta = \frac{1}{2C}$ gives $\mathrm{W_2H}(4C)$. □

Acknowledgements. It is my great pleasure to thank Prof. Li-Ming Wu for his warm encouragement. And I deeply appreciate the anonymous reviewer for his/her conscientious reading and many suggestions on the first version. This work is supported by NSFC (no. 11201456, no. 1143000182, no. 11371352), AMSS research grant (no. Y129161ZZ1), and Key Laboratory of Random Complex Structures and Data, Academy of Mathematics and Systems Science, Chinese Academy of Sciences (No. 2008DP173182).

References

[1] Bakry D., Barthe F., Cattiaux P., and Guillin A.: A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case, *Electron. Commun. Probab.* **13** (2008), 60-66

[2] Bakry D., Cattiaux P., and Guillin A.: Rates of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. *J. Funct. Anal.* **254** (2008), no. 3, 727-759

[3] Bakry D., Gentil I., and Ledoux M.: Analysis and geometry of Markov diffusion operators. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], **348**. Springer, Cham, 2014
[4] Barthe F., Cattiaux P., and Roberto C.: Interpolated inequalities between exponential and Gaussian, Orlicz hypercontractivity and application to isoperimetry. *Revista Math. Iberoamericana* **22** (2006), no. 3, 993-1067
[5] Bobkov S. G., Götze F.: Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.* **163** (1999), 1-28
[6] Cattiaux P., Guillin A.: Functional Inequalities via Lyapunov conditions. In *Optimal transportation, Theory and applications*, London Mathematical Society Lecture Notes Series, **413**, 274-287. Cambridge Univ. Press, 2014
[7] Cattiaux P., Guillin A., Wang F.-Y., Wu L.-M.: Lyapunov conditions for super Poincaré inequalities. *J. Funct. Anal.* **256** (2009), no. 6, 1821-1841
[8] Cattiaux P., Guillin A., and Wu L.-M.: A note on Talagrands transportation inequality and logarithmic Sobolev inequality, *Proba. Theory Relat. Fields* **148** (2010), no. **1-2**, 285-304
[9] Cattiaux P., Guillin A., and Zitt P.-A.: Poincaré inequalities and hitting times. *Ann. Inst. Henri Poincare Probab. Stat.* **49** (2013), no. 1, 95-118
[10] Djellout H., Guillin A., and Wu L.-M.: Transportation cost-information inequalities and applications to random dynamical systems and diffusions. *Ann. Probab.* **32** (2004), no. 3B, 2702-2732
[11] Evans L. C.: Partial differential equations. Second edition. Graduate Studies in Mathematics, **19**. American Mathematical Society, Providence, RI, 2010
[12] Gilbarg D., Trudinger N.: Elliptic partial differential equations of second order. Second edition. Grundlehren der Mathematischen Wissenschaften, **224**. Springer-Verlag, Berlin, 1983
[13] Gozlan N., Léonard C.: A large deviation approach to some transportation cost inequalities. *Probab. Theory Relat. Fields* **139** (2007), no. 1-2, 235-283
[14] Gozlan N., Léonard C.: Transport inequalities, a survey. *Markov Processes Relat. Fields* **16** (2010), 635-736
[15] Gozlan N., Roberto C., and Samson P. M.: A new characterization of Talagrands transport-entropy inequalities and applications. *Ann. Probab.* **39** (2011), no. 3, 857-880
[16] Gozlan N., Roberto C., and Samson P. M.: Hamilton Jacobi equations on metric spaces and transport entropy inequalities. (English summary) *Rev. Mat. Iberoam.* **30** (2014), no. 1, 133-163
[17] Guillin A., Léonard C., Wang F.-Y., and Wu L.-M.: Transportation information inequalities for Markov processes (II). arXiv:0902.2101
[18] Guillin A., Léonard C., Wu L.-M., and Yao N.: Transportation information inequalities for Markov processes, *Probab. Theory Relat. Fields* **144** (2009), no. 3-4, 669-696
[19] Holley R., Stroock D.: Logarithmic Sobolev inequalities and stochastic Ising models. *J. Stat. Phys.* **46** (1987), no. 5-6, 1159-1194
[20] Liu Y.: Gaussian integrability of distance function under the Lyapunov condition. *Electron. Commun. Probab.* **20** (2015), no. 9, 1-10
[21] Liu Y.: A link between the log-Sobolev inequality and Lyapunov condition. *Potential Anal.* **44** (2016), no. 4, 629-637
[22] Lott J., Villani C.: Hamilton-Jacobi semigroup on length spaces and applications. *J. Math. Pures Appl.* (9) **88** (2007), no. 3, 219-229
[23] Otto F., Villani C.: Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality, *J. Funct. Anal.* **173** (2000), no. 2, 361-400
[24] Talagrand M.: Transportation cost for gaussian and other product measures. *Geom. Funct. Anal.* **6** (1996), 587-600
[25] Villani C.: Topics in Optimal Transportation. Graduate Studies in Mathematics **58**, American Mathematical Society, Providence RI, 2003
[26] Villani C.: Optimal Transport: old and new. Grundlehren der Mathematischen Wissenschaften **338**, Springer-Verlag, Berlin, 2009

Yuan LIU, Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
E-mail address: `liuyuan@amss.ac.cn`